Effect of research participation versus usual clinical care in patients with rheumatic and musculoskeletal disorders: a prospective cohort study

Elisabeth Ginnerup-Nielsen, Robin Christensen, Henning Bliddal, Marius Henriksen

ABSTRACT

Objective To compare illness perception (IP), pain, functional level and health-related quality of life (HRQoL) between patients with musculoskeletal pain who participate versus those who do not participate in clinical research projects.

Methods Data were collected between 1 January 2019 and 31 December 2021 in patients visiting the Outpatient Osteoarthritis Clinic at Frederiksberg Hospital, Copenhagen, as part of either clinical research or regular treatment. Questionnaires were collected at baseline and after 10–18 months. Major outcome measure was the change from baseline to follow-up in the Brief Pain Inventory - Short Form (BPI-SF) item ‘Average pain’. Secondary outcome measures included the Brief Illness Perception Questionnaire (B-IPQ), measured only at baseline, the EuroQol (EQ-5D-3L), the Health Assessment Questionnaire Disability Index and PainDETECT.

Results 1495 patients were included with 358 (24%) categorised as research participants (exposed) and 1137 (76%) being non-participants (unexposed). The baseline B-IPQ item scores were generally more favourable in the exposed group with statistically significant standardised differences (SD) of 0.2–0.3. Similarly, an SD of 0.3 on the EQ-5D-3L score indicated a better HRQoL in the exposed group. At follow-up, 24% in the exposed group and 27% in the unexposed group, completed the questionnaires. The mean BPI-SF ‘Average pain’ difference between groups was −0.01 points (95% CI: −0.6 to 0.6). Similar clinically irrelevant differences were seen in the other outcomes.

Conclusions Among musculoskeletal pain patients, research participants report more positive IP and better HRQoL than non-participants. No additional effect of research participation was found in any outcome over time.

Trial registration number NCT03785561.

INTRODUCTION

Rheumatic and musculoskeletal disorders (RMDs) are some of the most common disorders in Europe, affecting around 20% of people over the age of 50 as well as being some of the highest contributors of years lived with disability.1 While pain is a common issue, symptoms, treatment options as well as treatment effects vary substantially.2 3 Illness perceptions (IP) are the thoughts an individual creates when faced with a health threat.4 These thoughts usually concern five coherent components; consequences, timeline, control/cure, identity and cause and eventually lead to a perception of the illness as either manageable or threatening. IP have been shown to be related to both pain, function and quality of life (QoL)5 6 as well as treatment choices and adherence7 8 in patients with musculoskeletal disorders. As people will tend to seek common sense between their perceptions and choices of management9 it seems reasonable to believe that certain IP traits will be prevalent in patients who choose to participate in research. As a hypothetical example, a patient perceiving her/his pain as untreatable may be less likely to engage in a clinical study while a patient with stronger
beliefs in the controllability of her pain, may be more likely to engage in research. As these perceptions have been shown to be predictive of changes in functional level over time,\(^\text{16}\) study participants could be both positively and negatively biased by means of their baseline perceptions.

Some studies have raised concerns regarding the external validity of especially randomised controlled trials as this method might attract a selected group of patients.\(^\text{11}\) No specific ‘trial effect’ has been identified when comparing study participants and patients eligible for the same study but not consenting to participate.\(^\text{12}\) Meanwhile, the often quite strict eligibility criteria and demands associated with participation in a clinical trial support the assumption that study participants share common characteristics, which do not necessarily represent the wider patient population. However, also observational designs are at risk of including data or surveying participants in a way that some members of the target population are less likely to be included or surveyed than others, known as ascertainment bias.\(^\text{13}\) This piece of bias has been noted within cardiac research for years with male participants being markedly more likely to be included in research than females\(^\text{14}\) and has recently been described within vaccine trials as well.\(^\text{15}\) An exaggerated effect (bias) of research participation could also be induced by participants being more compliant to treatments when followed in a study. This tendency has already to some extent been shown in earlier studies, known as the ‘Hawthorne effect’ described by Elton Mayo in 1946\(^\text{16}\) although the extent of this effect has been discussed.\(^\text{17}\) It seems likely that the extra attention study participants will normally gain from participation in a study might by itself evoke a feeling of safety or reassurance among patients. In line with this, in 2011, Brodersen et al\(^\text{18}\) found that women (n=362) living in an area where screening for breast cancer was offered, reported less negative perceptions concerning the risk of breast cancer compared with women living in an area where screening was not offered (n=568), even though they did not all attend the mammographies. Similarly, several studies have found an inexpedient, unrealistic optimism about benefitting personally from trial participation in phase one cancer trials among patients with terminal cancer,\(^\text{19, 20}\) indicating that study participants in this area are basically either selected from the beginning or optimism is induced by study participation.

From the previous research it seems fair to assume that participation in clinical research could be associated with better baseline levels which might be explained by research participants representing a highly selected phenotype already at recruitment. If so, the empirical impact and long-term outcome of such bias is unclear. The objectives of this study were to compare the baseline IPs among research participants with those of non-participants and to assess the longitudinal effect of research participation on frequently used patient-reported outcome measures in patients with RMDs.

**METHODS**

**Study design**

This was a prospective cohort study conducted according to a published protocol (www.clinicaltrials.gov) and reported according to Strengthening the Reporting of Observational Studies in Epidemiology guidelines\(^\text{21}\) (online supplemental file 1). The study was carried out at the Outpatient Osteoarthritis Clinic at Bispebjerg-Frederiksberg Hospital, Copenhagen, Denmark, between 1 January 2019 and 31 December 2021. Last participant was included on 28 February 2021 to allow for at least 10 months follow-up. Participants were recruited among regular patients or study participants at the clinic in this time span. The exposed group was defined as participants being enrolled in a health research study with at least one study related visit during the study period (until 28 February 2021). We defined a research study as any study that had been approved by the local health research ethics committee. Under Danish legislation, studies that involve any kind of pharmacological or non-pharmacological experiment on human beings require health research ethical approval while studies only including questionnaires or interviews are exempt from health research ethical approval.

Patients visiting the clinic as part of a regular health assessment or treatment, not being part of any research study were considered unexposed. Patients, who were enrolled in a research study between the two visits (eg, had at least one visit as part of a research study) were considered part of the ‘exposed’ group. No power calculations were done due to the exploratory nature of the study, but we expected around 1150 yearly visits at the clinic with almost all patients being eligible for this study.

**Participants and data**

 Eligible participants were all patients \(\geq 18\) years visiting the osteoarthritis outpatient clinic due to an RMD, consenting to participate and being able to understand Danish. Data were collected via touch screens during each patient’s visit at the clinic. All patients visiting the clinic (whether participating in research or not) filled out questionnaires as part of daily practice and consented that these could potentially be used by other health professionals. Hence, we were able to collect data from patients that were currently not research participants and did not consider themselves as such.

**Objectives and outcomes**

Our primary objective was to compare the change in pain over time assessed by the primary outcome of change from baseline in the Brief Pain Inventory - Short Form (BPI-SF) item ‘average pain’.

The secondary objective was to explore baseline group differences in IP based on the Brief Illness Perception Questionnaire (B-IPQ). To further describe our exposed and unexposed groups in terms of pain patterns all patients responded to the PainDETECT, EQ-5D-3L and
the Health Assessment Questionnaire Disability Index (HAQ-DI) collected at baseline.

Secondary outcome measures included change from baseline in the BPI-SF items ‘least pain’, ‘worst pain’ and ‘pain right now’ as well as a composite score of the seven functional items of the BPI-SF defined as ‘pain interference’. Further secondary outcome measures included changes from baseline in the EuroQol (EQ-5D-3L) and the HAQ-DI.

**The Brief Illness Perception Questionnaire**

A generic questionnaire developed to measure IP in a variety of illnesses. The questionnaire is patient-reported and assesses perceptions on the following five dimensions: Identity, Cause, Timeline, Consequences and Cure-Control. The Danish version contains eight items scored on a numerical 1–10-point rating scale and a memo field allowing patients to describe perceived causes of their condition. In items 1, 2, 5, 6 and 8 a higher score indicates a more threatening view of the illness (eg, item 2: How long do you think your illness will continue?) while a higher score in items 3, 4 and 7 indicates a less threatening view (eg, item 3: How much control do you feel you have over your illness?). As proposed in the literature, we replaced the word ‘illness’ with ‘pain’, given that all our patients had musculoskeletal pain in differing sites. The B-IPQ has been validated in populations with acute and chronic low back pain. An acceptable inter-item consistency of 0.73 (Cronbach’s α; 95% CI: 0.67 to 0.83) has been found and moderate overall test-retest reliability with acceptable intraclass correlation coefficient ranging from 0.65 to 0.88 for the B-IPQ items. Minimal detectable changes on a group level of between 0.40 and 0.56 points have been described.

No minimal clinically important difference (MCID) on each item has been defined.

**The Brief Pain Inventory**

It was previously known as the BPI is a self-administered questionnaire that was originally designed to assess cancer pain. It is now also used and validated as a generic pain questionnaire for other chronic pain conditions and is available in a long form (15 questions) and a short form (9 questions). The BPI measures the intensity of pain experienced within the last 24 hours and how much pain has interfered with seven daily activities, including general activity, walking, work, mood, enjoyment of life, relations with others and sleep. In this study we used the short version (BPI-SF). The BPI-SF is based on numerical rating scales (0–10) rating pain severity and pain interference in daily activities. The arithmetic mean of the four pain severity items can be used as measures of pain severity, but within research, the two items ‘Average Pain’ and ‘Worst pain’ are often reported as main single item outcomes. Furthermore, the authors of the questionnaire recommend that a composite score of how pain interferes with seven daily activities is reported. A lower score indicates less pain interference. Although not officially recognised, the MCID on both the BPI—long and short form is considered to be 1–2 points. A Danish version of the BPI-SF is available on www.manderson.org.

**The EuroQol (EQ-5D-3L)**

A self-reported outcome measure that provides a simple measure of health-related quality of life (HR-QoL). It is applicable to a wide range of health conditions. The EQ-5D-3L consists of the EQ-5D descriptive system and the EQ Visual Analogue Scale (EQ VAS). We did not use the EQ VAS for this study. The EQ-5D-3L descriptive system comprises the following five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension has three levels: ‘no problems’, ‘some problems’ and ‘extreme problems’. Scores on the five dimensions are recalculated to an index between −0.624 corresponding to being seriously affected and 1.000 corresponding to not being affected at all. Minimally important difference of the EQ-5D in musculoskeletal disorders is around 0.03–0.54.

**The Health Assessment Questionnaire Disability Index**

It measures self-reported ability in different domains. HAQ-DI contains eight sections: dressing, arising, eating, walking, hygiene, reach, grip and activities with two or three questions for each section. Each section can be scored from 0 (without any difficulty) to 3 (unable to do). For each section the score given to that section is the worst score within the section, that is, if one question is scored 1 and another 2, then the score for the section is 2. In addition, if an aid or device is used or if help is required from another individual, then the minimum score for that section is 2. The 8 scores of the eight sections are summed and divided by 8. If one section is not completed by a subject, then the summed score would be divided by 7. Scores of 0–1 are generally considered to represent mild-to-moderate difficulty, 1–2 moderate-to-severe disability and 2–3 severe-to-very severe disability. The Danish adaptation of the HAQ-DI has been validated on patients with rheumatoid arthritis showing good psychometric properties with an excellent test-retest reliability Cronbach’s alpha coefficient of >0.9, and the ability to validly distinguish between patient-perceived severity and functional status. The MCID is around 0.22 points.

**PainDETECT Questionnaire (PDQ)**: A self-reported measure developed to detect neuropathic pain components in adult patients. It consists of seven questions addressing the quality of neuropathic pain symptoms scored from 0 to 5 (never=0, hardly noticed=1, slightly=2, moderately=3, strongly=4, very strongly=5), four questions addressing the patterns of pain (scored from −1 to +1) and one question concerning radiation of pain into other parts of the body (yes (+2)/no (0)). The questionnaire also includes a full body drawing used for diagnostics enabling a visual description of painful body parts. A total score ranging from −1 to 38 is calculated based on an algorithm with pain intensity ratings not included in this score. The final score between −1 and 38 indicates...
the likelihood of a neuropathic pain component. A score of ≤12 indicates that pain is unlikely to have a neuropathic component, while a score of ≥19 suggests that pain is likely to have a neuropathic component. The questionnaire has been found to have an adequate internal consistency (Cronbach’s alpha=0.83) and a specificity as well as sensibility of 84%. A Danish version of the PDQ is available.

Statistical analysis
All analyses were prespecified in a statistical analysis plan (online supplemental file 2). Participant characteristics and baseline measures including the B-IPQ item scores for the analysis population were summarised and presented separately for the two groups (referred to as exposed and unexposed). Continuous data were summarised by means and SDs and categorical data by numbers and percentages. The group distributions of observed baseline covariates were compared using standardised differences. Differences in primary and secondary outcome measures over time are presented in standardised differences and 95% CI at baseline and means and 95% CI at follow-up.

Due to the exploratory nature of this study, we did not apply explicit adjustments for multiplicity, but prespecified that we would interpret the multiple analyses in a prioritised order with analyses performed in sequence until eventually one of the analyses failed to show statistically significant difference (level of statistical significance: 0.05). Both primary and secondary objectives were analysed based on a superiority framework; reporting both crude (unadjusted) estimates and analyses adjusted for potential pre-exposure covariates. These included: the primary analysis with adjustment for (1) sex and baseline age and (2) sex, baseline age and outcome levels at baseline.

Due to the significant amount of attrition, we decided not to perform sensitivity analyses based on multiple imputations for repeated replacement of missing data at year 1.

Results were reported as least squares means with the corresponding SEs derived from the mixed models, while differences between groups were reported with 95% CIs.

Analyses were performed using SAS Studio V.9.4.

RESULTS
Between 01 January 2019 and 28 February 2021, 1896 patients were registered in the outpatient osteoarthritis clinic. Of these, 12 (0.6%) did not consent to share their data, 378 (20%) did not respond to the BPI-SF item ‘Average pain’ at baseline and 11 (0.6%) were not having a musculoskeletal disorder. Hence, 1495 patients were included in the study with 358 (24%) being exposed and 1137 (76%) being unexposed. A median follow-up time of 391 days from baseline was seen. In the exposed group, 118 (33%) had a follow-up visit while 86 (24%) had responded to the BPI-SF item ‘Average pain’. In the unexposed group, 306 (27%) had a follow-up visit and all had responded to the BPI-SF ‘Average pain’ at the visit (figure 1).

All research studies conducted in the outpatient clinic during data collection were studies requiring health research ethical approval.

Baseline characteristics are summarised in table 1. Small to moderate group differences were seen in all demographic variables, with relatively fewer women in the exposed group (SD: −0.29; (p≤0.0001)), markedly more patients having osteoarthritis and markedly less patients having inflammatory arthritis in the unexposed group. Also, patients having an undefined ‘other’ diagnosis, which was usually unspecific pain at different sites, were only seen in the unexposed group. A mean PDQ score of 10.0 (SD 6.7) in the unexposed group and 10.8 (SD 7.4) in the exposed group indicated that both groups were unlikely to have neuropathic pain.

Figure 1 Flowchart illustrating participant flow. BPI, Brief Pain Inventory.
Primary analyses

In the primary analysis based on crude changes in the intention to survey population, neither statistically nor clinically relevant group difference in change from baseline was observed at 1-year follow-up in the primary outcome BPI-SF item ‘average pain’ (group difference −0.01 points (95% CI: −0.6 to 0.6)) (Table 2).

As illustrated in Table 1, baseline differences in B-IPQ scores were small but statistically significant, with standardised differences between −0.3 points (p=0.0001) for item Consequences and 0.1 (p≤0.05) for item Treatment Control were seen with the unexposed group generally perceiving their pain as a greater threat to them than the exposed group with higher scores on the items 1, 5, 6 and 8 and lower scores on items 3 and 4. Only item 2 deviated with participants in the exposed group expecting their pain to last for a longer time than participants in the unexposed group. No group difference was seen in item 7. Similar small baseline differences were seen in the other outcome variables with the exposed group consistently showing more favourable scores however, the EQ-5D was the only variable reaching a clear clinically
relevant difference (0.047 index points; standardised difference: 0.3, p=0.0001).

**Secondary objectives**

Similar to the primary outcome measure, neither statistically nor clinically relevant group differences in changes from baseline were observed at 1-year follow-up in the secondary outcomes BPI-SF items: ‘worst pain’, ‘least pain’, ‘pain right now’ and ‘pain interference’, HAQ-DI or EQ-5D (table 2). The sensitivity analyses of the primary outcome BPI-SF ‘average pain’ including changes adjusted for (1) sex and age, and (2) sex, age and level at baseline confirmed the primary analysis with clinically irrelevant and statistically insignificant differences in changes of respectively 0.003 points (95% CI: −0.6 to 0.6) and −0.009 points (95% CI: −0.5 to 0.5). Similar small and irrelevant differences in changes were observed for the secondary outcomes (table 2).

**DISCUSSION**

In this cohort study based on 1495 patients with musculoskeletal pain, we found small but consistent baseline

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Changes from baseline to follow-up in primary and secondary outcome measures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unexposed</td>
</tr>
<tr>
<td></td>
<td>Mean (SE)</td>
</tr>
<tr>
<td><strong>Crude (unadjusted) change from baseline in</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Primary outcome:</strong></td>
<td></td>
</tr>
<tr>
<td>BPI-SF item ‘average pain’</td>
<td>0.03 (0.1)</td>
</tr>
<tr>
<td><strong>Secondary outcomes:</strong></td>
<td></td>
</tr>
<tr>
<td>BPI-SF item ‘worst pain’</td>
<td>0.2 (0.2)</td>
</tr>
<tr>
<td>BPI-SF item ‘least pain’</td>
<td>0.4 (0.1)</td>
</tr>
<tr>
<td>BPI-SF item ‘pain right now’</td>
<td>0.2 (0.2)</td>
</tr>
<tr>
<td>BPI-SF item ‘pain interference’</td>
<td>0.1 (0.1)</td>
</tr>
<tr>
<td>HAQ</td>
<td>0.03 (0.02)</td>
</tr>
<tr>
<td>EQ-5D-3L</td>
<td>0.002 (0.010)</td>
</tr>
</tbody>
</table>

**Analyses adjusted for sex and age**

<table>
<thead>
<tr>
<th></th>
<th>Unexposed</th>
<th>Exposed</th>
<th>Difference (95% CI)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SE)</td>
<td>Mean (SE)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Primary outcome:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BPI-SF item ‘average pain’</td>
<td>0.03 (0.1)</td>
<td>0.04 (0.3)</td>
<td>0.003 (−0.6 to 0.6)</td>
<td>0.93</td>
</tr>
<tr>
<td><strong>Secondary outcomes:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BPI-SF item ‘worst pain’</td>
<td>0.2 (0.2)</td>
<td>0.4 (0.3)</td>
<td>0.2 (−0.4 to 0.9)</td>
<td>0.50</td>
</tr>
<tr>
<td>BPI-SF item ‘least pain’</td>
<td>0.4 (0.1)</td>
<td>0.6 (0.3)</td>
<td>0.2 (−0.4 to 0.8)</td>
<td>0.62</td>
</tr>
<tr>
<td>BPI-SF item ‘pain right now’</td>
<td>0.2 (0.2)</td>
<td>0.8 (0.3)</td>
<td>0.5 (−0.1 to 1.2)</td>
<td>0.12</td>
</tr>
<tr>
<td>BPI-SF item ‘pain interference’</td>
<td>-0.1 (0.1)</td>
<td>0.1 (0.2)</td>
<td>0.2 (−0.3 to 0.7)</td>
<td>0.47</td>
</tr>
<tr>
<td>HAQ</td>
<td>0.04 (0.02)</td>
<td>0.02 (0.03)</td>
<td>-0.02 (−0.1 to 0.1)</td>
<td>0.69</td>
</tr>
<tr>
<td>EQ-5D-3L</td>
<td>0.001 (0.010)</td>
<td>0.020 (0.018)</td>
<td>0.019 (−0.020 to 0.059)</td>
<td>0.33</td>
</tr>
</tbody>
</table>

**Analyses adjusted for sex, age and level at baseline**

<table>
<thead>
<tr>
<th></th>
<th>Unexposed</th>
<th>Exposed</th>
<th>Difference (95% CI)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SE)</td>
<td>Mean (SE)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Primary outcome:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BPI-SF item ‘average pain’</td>
<td>0.01 (0.1)</td>
<td>0.001 (0.2)</td>
<td>-0.009 (−0.5 to 0.5)</td>
<td>0.97</td>
</tr>
<tr>
<td><strong>Secondary outcomes:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BPI-SF item ‘worst pain’</td>
<td>0.1 (0.1)</td>
<td>0.3 (0.3)</td>
<td>0.2 (−0.4 to 0.7)</td>
<td>0.54</td>
</tr>
<tr>
<td>BPI-SF item ‘least pain’</td>
<td>0.4 (0.1)</td>
<td>0.4 (0.2)</td>
<td>0.01 (−0.5 to 0.5)</td>
<td>0.96</td>
</tr>
<tr>
<td>BPI-SF item ‘pain right now’</td>
<td>0.2 (0.1)</td>
<td>0.6 (0.3)</td>
<td>0.3 (−0.2 to 0.9)</td>
<td>0.24</td>
</tr>
<tr>
<td>BPI-SF item ‘pain interference’</td>
<td>0.1 (0.1)</td>
<td>0.1 (0.2)</td>
<td>0.1 (−0.4 to 0.6)</td>
<td>0.72</td>
</tr>
<tr>
<td>HAQ</td>
<td>0.03 (0.02)</td>
<td>0.02 (0.03)</td>
<td>-0.01 (−0.1 to 0.1)</td>
<td>0.84</td>
</tr>
<tr>
<td>EQ-5D-3L</td>
<td>0.006 (0.009)</td>
<td>0.015 (0.015)</td>
<td>0.009 (−0.024 to 0.043)</td>
<td>0.58</td>
</tr>
</tbody>
</table>

Values are least squares means (SEs) unless otherwise stated. Brief Illness Perception Questionnaire: a higher score reflects a more threatening view of the illness in item 1, 2, 5, 6 and 8, and a less threatening view in item 3, 4 and 7. BPI-SF, Brief Pain Inventory - Short Form: a lower score reflects less pain interference; EQ-5D-3L, a higher score reflects a better quality of life; HAQ-DI, The Health Assessment Questionnaire Disability Index (DI): a lower score reflects less disability.
differences across all outcomes with the research participants (exposed group) generally reporting less pain interference, less disability and better HR-QoL with group difference on the EQ-5D-3L being clinically relevant, with a difference of 0.047 index points corresponding to a (standardised difference of 0.3, \( p=0.0001 \)).

These results were reflected in the B-IPQ with standardised group differences on the B-IPQ items of between 0.2 and 0.3, with the exposed group perceiving their pain as less threatening compared with the unexposed group except on B-IPQ item 2 _Timeline_ where the exposed group perceived their pain to be more chronic than the unexposed group and item 7 representing illness _Coherence_ where no group difference was seen. At follow-up, no group differences in changes were seen in either pain function or HR-QoL.

These results are similar to what we have earlier found in a cross-sectional cluster analysis on 1552 people with knee pain,\(^6\) where participants who reported lower knee pain levels and higher HR-QoL also perceived their knee pain as less threatening compared with people with higher pain levels and lower HR-QoL. Like the present results, there were no group differences in item 7 concerning illness _Coherence_, suggesting that this item may not relate to the rest of the B-IPQ items.

While diagnoses in our population were more benign, the results were in accordance with what has been found in qualitative studies of oncology patients and seem to reflect a fundamental trait in humans having an initial face of optimism when entering a trial. Ongoing participation would not seem to impact QoL to a similar extent, possibly reduced by time-consuming trial schedules.\(^{39,40}\)

The higher HR-QoL scores may have been influenced by differences in the distribution of the patient population with 78.3% in the unexposed group having OA but only 67.6% in the exposed group, with a larger proportion of inflammatory arthritis. As patients with rheumatoid arthritis have been found to evaluate their HR-QoL as lower than patients with osteoarthritis,\(^{41}\) the baseline differences we observed could simply be derived from differences in the baseline distribution of patient populations. Meanwhile, this difference has been questioned\(^{42}\) and it seems reasonable to hypothesise that the higher HR-QoL score simply reflects that patients with more mental surplus will more likely engage in research.

We found no effect of being exposed on changes in any measures in our study population. There was, however, a considerable limitation with regards to this result as only one in four participants were seen at a follow-up visit as defined in the study. Thus, the main part of participants has been lost during this time either due to a positive treatment effect or the contrary, that is, disappointment with the result.

Vist _et al_, in a Cochrane review from 2008\(^{12}\) explored a possible trial effect, in 85 trials (mostly non-randomised) with over 86640 patients with differing diseases. They suggested that ‘participating in a randomized controlled trial is likely to result in similar outcomes to having similar treatment outside of the trial’. Although only a few of the included studies concerned musculoskeletal pain disorders, pain was a regular outcome and our results are in accordance with this conclusion. Of note, the control groups in a major part of the included trials in Vist _et al_,\(^{12}\) represented eligible patients who refused randomisation or had treatment preferences. Hence, the representativeness of randomised trial samples and the risk of conducting ascertainment bias when recruiting for trials may likely still be an issue, which is also reflected in recent research.\(^{43}\) In the present study we do not know whether patients in the unexposed group were eligible for a trial, but as the patient populations in the two groups differ to some extent it is likely not the case for most of the patients.

Participation in a clinical study may affect participants, not only by the actual intervention, but also by contextual factors related to the interventional setting. A few studies have explored the effect of being treated in an institution where participation in health research is common. Two register based studies exploring the effect of research participation on mortality rates in the British National Health Service Trusts\(^{44,45}\) found the number of citations correlated with lower mortality rates and a systematic review of 13 trials\(^{46}\) suggested that adherence to guidelines is optimised and doctors involved in trials tend to prescribe treatments earlier than doctors in routine practice. As both our exposed and unexposed group were treated in the outpatient osteoarthritis clinic, an institution with comprehensive research activity, the lack of group difference could be due to the unexposed group being exposed to a similar setting as the exposed group. Meanwhile, in 2002, Majumdar _et al_\(^{47}\) questioned the effect of research activity as they found that American sites that had taken part in a trial demonstrating that ACE inhibitors were beneficial following myocardial infarction, were no more likely to adopt their use in this group of patients than sites who had not participated in the trial.

Although it seems rational to hypothesise that research active sites are more likely to follow guidelines and thereby offer optimised treatments, the observed differences have been small and the effect of other contextual factors such as research traditions and organisation of healthcare systems may be of greater importance than the number of citations. This assumption is supported by a Danish retrospective cohort study measuring the effects of trial participation on medication prescription among physicians.\(^{48}\) Based on 10 general practices participating in industry-initiated trials on asthma drugs (N=5439) and 165 control practices not participating in any trials (N=59574), no difference was found in adherence to guidelines, but trial active practices were more likely to prescribe drugs from the trial sponsor than control practices.

Although the baseline differences in HR-QoL and IP among study participants and non-participants should be interpreted with caution, they support the notion of
patients participating in research being a selected group. To augment the external validity of trials especially in acute settings, some have used retrospective consent eventually minimising a trial effect, and potentially also baseline differences, hypothesising that patients with more comorbidities and less mental surplus could be more reluctant to immediately accept study participation, but would accept retrospectively. When exploring objective outcomes, hybrid trials with endpoints collected through routine non-standardised healthcare visits may augment the generalisability of the results. Finally, conducting pragmatic trials in real world settings with less strict inclusion criteria as well as being more proactive in the recruitment process may also allow for inclusion of a more diverse population.

This study has limitations, first and foremost, the large dropout rates of 76% in the exposed group and 73% in the unexposed group are—although similar in both groups reducing confidence in the follow-up data. Furthermore, we defined our exposed group as patients who were currently participating in a research study. We cannot know whether people in the unexposed group have been participating in trials before, but we do not expect earlier trial participation to impact on actual levels of pain, function and HR-QoL per se. The baseline group imbalance of conditions is, although relatively limited in absolute numbers, a limitation. Furthermore, the unexposed group may include both patients who were eligible to a study but chose not to participate and some who were ineligible due to, for example, comorbidities. These patients may perceive their disease differently and differ in terms of health-related outcomes, but as we included many different studies with very different inclusion criteria, we were not able to explore these differences.

Finally, we only included patients from one site limiting the generalisability of the study.

In conclusion, these data suggest that patients with musculoskeletal pain enrolling in clinical research have a more positive IP and better HR-QoL than patients that do not participate in research studies. No effect of research participation was identified in either musculoskeletal pain, functional level or HR-QoL over time.

Acknowledgements The authors would like to thank data managers Christian Cato Holm and Peter Krusager for all their help with the management of the data.

Contributors EG-N, MH, HB and RC conceptualised and designed the study. EG-N and HB substantially contributed to the acquisition of clinical data. EG-N and RC processed the data and performed the statistical analysis. All authors contributed to interpretation of the data. EG-N wrote the article, and all coauthors revised the manuscript critically for important intellectual content. EG-N is responsible for the overall content as a guarantor and controlled the decision to publish. All authors agreed on the final content of the submitted manuscript.

Funding This study received no specific funding, but The Parker Institute, Bispebjerg and Frederiksberg Hospital is supported by a core grant from the Oak Foundation (OCAY-N-347-OML).

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval A request was made to the Capital Region Research Ethics Committee with journal number: 18016730. According to Danish laws this kind of research (only based on questionnaires) need no approval. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. Access to data may be obtained upon reasonable request by contacting the corresponding author.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iD Elisabeth Ginnerup-Nielsen http://orcid.org/0000-0001-6314-769X

REFERENCES


47 Majumdar SR, Chang W-C, Armstrong PW. Do the investigative sites that take part in a positive clinical trial translate that evidence into practice *The American Journal of Medicine* 2002;113:140–5.

